



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2013

---

## **An electrophysiological study of print processing in kindergarten: the contribution of the visual n1 as a predictor of reading outcome**

Brem, Silvia ; Bach, Silvia ; Kujala, Janne V ; Maurer, Urs ; Lyytinen, Heikki ; Richardson, Ulla ; Brandeis, Daniel

**Abstract:** Sensitivity to print is characterized by a left occipito-temporal negativity to words in the event-related potential N1. This sensitivity is modulated by reading skills and may thus represent a neural marker of reading competence. Here we studied the development of the N1 in regular and poor readers from preschool age to school age to test whether the amplitude of the N1 predicts children's reading outcomes. Our results suggest a predictive value of the print-sensitive negativity over the right hemisphere. Whether this N1 may serve as a biomarker to improve prognosis in preliterate children should be clarified in future studies.

DOI: <https://doi.org/10.1080/87565641.2013.828729>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-87937>

Journal Article

Accepted Version

Originally published at:

Brem, Silvia; Bach, Silvia; Kujala, Janne V; Maurer, Urs; Lyytinen, Heikki; Richardson, Ulla; Brandeis, Daniel (2013). An electrophysiological study of print processing in kindergarten: the contribution of the visual n1 as a predictor of reading outcome. *Developmental Neuropsychology*, 38(8):567-594.

DOI: <https://doi.org/10.1080/87565641.2013.828729>

## **An Electrophysiological Study of Print Processing in Kindergarten: The Contribution of the Visual N1 as a Predictor of Reading Outcome**

**Running head:** *The visual N1 as a predictor of reading outcome?*

### **Authors:**

Silvia Brem<sup>1,2</sup>, Silvia Bach<sup>1,2,7</sup>, Janne V. Kujala<sup>1,8</sup>, Urs Maurer<sup>2,4</sup>, Heikki Lyytinen<sup>3</sup>, Ulla Richardson<sup>1\*</sup>, Daniel Brandeis<sup>2,5,6,7\*</sup>

*\* the second last and last authors have equally contributed to this article*

- 1) Agora Center, University of Jyväskylä, Finland
- 2) Department of Child and Adolescent Psychiatry, University of Zurich, Switzerland
- 3) Agora Center, University of Jyväskylä, Jyväskylä, Finland, and Department of Psychology, University of Jyväskylä, Jyväskylä, Finland
- 4) Department of Psychology, University of Zürich, Switzerland
- 5) Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland
- 6) Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim/ Heidelberg University, Mannheim, Germany
- 7) Neuroscience Center Zurich, University of Zurich and ETH Zurich, Switzerland
- 8) Department of Mathematical Information Technology, University of Jyväskylä, Finland.

### **Corresponding author:**

Silvia Brem

Department of Child and Adolescent Psychiatry, University of Zürich

Neumünsterallee 9, CH- 8032 Zürich

[sbrem@kjp.d.uzh.ch](mailto:sbrem@kjp.d.uzh.ch), phone: +41 43 499 2760

**Running head: *The Visual N1 as a predictor of reading outcome?***

**Abstract**

Sensitivity to print is characterized by a left occipito-temporal negativity to words in the event-related potential N1. This sensitivity is modulated by reading skills and may thus represent a neural marker of reading competence. Here we studied the development of the N1 in regular and poor readers from preschool age to school age to test whether the amplitude of the N1 predicts children's reading outcomes.

Our results suggest a predictive value of the print-sensitive negativity over the right hemisphere. Whether this N1 may serve as a biomarker to improve prognosis in preliterate children should be clarified in future studies.

**Introduction**

An accurate prediction of reading and spelling outcomes before school enrollment would facilitate supporting those children with a poor prognosis, because customized trainings could be offered before any reading and spelling problems emerge. At present, difficulties in reading and/or spelling, such as those seen in children with developmental dyslexia, can only be diagnosed one to two years after the start of reading instruction. Early intervention would not only be most beneficial from a didactic point of view, but could also prevent associated emotional and behavioral problems (Willcutt & Pennington 2000; Mugnaini, Lassi, La Malfa, G. & Albertini, 2009). A major goal therefore is to improve prediction of reading and spelling difficulties at preschool age.

The body of research seeking reliable predictors of reading and spelling skills at preschool age is constantly growing. Among the most important behavioral predictors indicating later reading success are phonological and phonemic awareness, letter knowledge, and rapid automatized naming (RAN) of pictures, digits, colors, or letters (Bowey, 1995; Catts, Fey, Zhang, & Tomblin, 2001; de Jong & van der Leij, 1999; Goswami et al., 2011; Juel, 1986;

Lyytinen et al., 2004; Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007; Puolakanaho et al., 2007; Scarborough, 1990; Wagner & Torgesen, 1987; Wagner et al., 1997; Wimmer, Landerl, Linortner, & Hummer, 1991; Wolf, 1986). However, a child's literacy environment, socio-economic status (SES), and family reading history also influence his or her development of literacy skills (Bowey, 1995; Juel, 1986). In contrast to the considerable number of studies that examine behavioral or demographic measures that may predict later success in reading and/or spelling, researchers have neglected to examine other potential predictors such as brain measures as thoroughly. Due to the easy application of electroencephalography (EEG) in young children or even infants, most research has focused on electrophysiological responses and their contribution to prediction. EEG is a non-invasive technique with an excellent temporal resolution in the millisecond time range and is less prone to artifacts generated by motion of the head or body, as compared to other imaging techniques such as magnetic resonance imaging (MRI). The recording of the EEG allows to analyze event-related potentials (ERPs) that are time-locked neurophysiological responses to specific—e.g., sensory—events. In dyslexia research, auditory event-related potentials have demonstrated impressive predictive value concerning the development of reading skills; for instance, ERPs to syllable sounds recorded in newborns allowed for accurate predictions to be made of those children's language and verbal memory skills 2.5 to 6.5 years later (Guttorm, Leppanen, Hamalainen, Eklund, & Lyytinen, 2010; Guttorm et al., 2005), while ERPs to non-speech pitch sound discrimination differentiated between children with and without familial risk for developmental dyslexia (Leppanen et al., 2010). In newborns, they were further also associated with children's phonological skills at the age of 3.5, letter knowledge at the age of five, and reading skills at the age of nine (Leppanen et al., 2010). Amazingly, auditory ERPs to speech and nonspeech sounds recorded in infants even allowed to discriminate between infants classified as children with dyslexia, children with poor, or children with normal reading skills eight years later (Molfese, 2000). In addition, in somewhat older children at preschool age, the ERPs evoked by automatic phoneme processing substantially contributed to prognosis; the auditory response not only improved

the prediction of reading outcome over behavioral measures alone, but it was also the only measure capable of predicting long-term reading outcomes in fifth grade (Maurer et al., 2009). Recently, the contributions of structural and functional MRI (fMRI) in predicting reading skills were examined and yielded promising results as well. Both specific patterns of functional activation during phonological processing and the morphology of the white and grey matter at the beginning of a school year contributed to the prediction of decoding skills later in the same school year: In more detail, differences in the functional activation of the left middle temporal, the right fusiform, the right middle occipital and the middle frontal gyri along with differences in grey matter density in the right fusiform gyrus and white matter density in left superior temporal and inferior parietal regions predicted later decoding ability (Hoeft et al., 2007). Furthermore, the combination of the right prefrontal activation in a rhyming task and the white matter organization in the left superior longitudinal fasciculus informed about the long-term reading gains of children with developmental dyslexia (Hoeft et al., 2011). Lately, it has been shown that structural as well as functional (Raschle, Zuk, & Gaab, 2012) differences related to dyslexia are already detectable in preschool children, before they learn to read (Raschle, Chang, & Gaab, 2010). In children with a high familial risk of developing reading problems, a reduced amount of grey matter was found in the bilateral parieto-temporal and left occipito-temporal regions (Raschle et al., 2010). Structural changes are thus present before children learn to read and may therefore be used as markers to complement preschool predictions of children's reading outcomes.

Another neural correlate that may be well suited to inform about children's reading outcomes at preschool age is the sensitivity of the occipito-temporal cortex to print. A print-sensitive response in the ERP and fMRI can be measured when comparing the visual processing of print (words, pseudowords) to falsefonts or symbol strings. Print sensitivity may emerge because specific neuronal populations within the ventral occipito-temporal cortex are selectively tuned to print during reading acquisition (Dehaene & Cohen, 2007; Dehaene, Cohen, Sigman, & Vinckier, 2005). Alternatively, print-sensitive responses can be explained by an interactive account, assuming that the left ventral occipital cortex serves as an

interface, integrating bottom-up visual information and top-down predictions from phonological and semantic areas that are based on learning and prior experience (Price & Devlin, 2011). In the ERP, a first print-sensitive response appears as a characteristic, left occipito-temporal negativity (N1 N170), which is more pronounced to printed words as compared to symbol strings in normal reading children (Brem et al., 2009; Maurer et al., 2006; Maurer et al., 2011) and adults (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Brem et al., 2005; Mahé, Bonnefond, Gavens, Dufour, & Doignon-Camus, 2012; Schendan, Ganis, & Kutas, 1998). This print sensitive N1 peaks at around 150-250ms after stimulus onset. Estimations of source locations (Brem et al., 2006; Maurer, Brem, Bucher, & Brandeis, 2005; Proverbio, Zani, & Adorni, 2008), intracranial recordings (Allison, McCarthy, Nobre, Puce, & Belger, 1994), and functional magnetic resonance imaging (fMRI) results (Baker et al., 2007; Brem et al., 2006; Cohen et al., 2002; Kronschnabel, Schmid, Maurer & Brandeis, 2013; Vinckier et al., 2007) converge in showing that this print sensitive response originates in the basal left occipito-temporal lobe, which is also referred to as the visual word form system (VWFS). Print sensitivity, reflected by the N1, is usually absent in non-reading preschool children. It develops when children learn grapheme-phoneme correspondences, the basic principle of alphabetic languages (Brem et al., 2010) and becomes most pronounced in young beginning readers (Maurer et al., 2005; Maurer et al., 2006; Maurer et al., 2011; Parviainen, Helenius, Poskiparta, Niemi, & Salmelin, 2006). In adolescence and adulthood, the print-sensitive left hemispheric negativity is often reduced again (Brem et al., 2006; Brem et al., 2009). The quality of the print-sensitive response within the N1 window is also modulated by reading skills. Adults with dyslexia showed a clearly reduced letter-string-specific response in the left occipito-temporal cortex as compared to normal readers when using an explicit word-processing task in a magnetoencephalography study (Helenius, Tarkiainen, Cornelissen, Hansen, & Salmelin, 1999). In addition, second-grade children (Maurer et al., 2007) or pre-adolescent children (Araujo, Bramao, Faisca, Petersson, & Reis, 2012) with developmental dyslexia exhibited a reduced print-sensitive N1 in the ERP when performing implicit print-processing tasks. Interestingly, when the children with dyslexia in

Maurer and colleague's (2011) study were tested again three years later in fifth grade, no attenuation in their N1 print sensitivity was observed (Maurer et al., 2011). Even though print sensitivity in the N1 seems to develop when children learn to read, some differences in the N1 distribution between words and symbol strings can already be seen in particular groups of non-reading preschool children. A more pronounced negativity to words than symbols strings over the right hemisphere was revealed in preschool children with high letter knowledge within the N1 interval. This right lateralized negativity was delayed in children with low letter knowledge and started only after the N1 window at around 430ms. These findings thus suggest that a critical degree of early literacy through familiarity with print is necessary to allow for rapid, automatized processing of letter strings (Maurer et al., 2005). A retrospective analysis of the same children, however, revealed a marginal right lateralized print-sensitive negativity within the N1 component also among preschool children exhibiting poor reading skills at school age (Maurer et al., 2007). This finding questions the meaning of the right lateralized pre-literate print-sensitive N1 negativity for the process of learning to read. To summarize, the visual, print sensitive N1 develops during literacy acquisition, shows an inverted U-shaped maturation curve and is associated with reading skills.

Despite some evidence from structural MRI (Raschle et al., 2010) and ERP (Bach, Richardson, Brandeis, Martin, & Brem, 2013) studies that the activity of visual print-processing areas may contribute to the prediction of reading outcome, only one study to date has examined the potential of the print-sensitive response in the N1 and the corresponding VWFS activation to predict reading outcomes. This study showed that the inclusion of the print-sensitive fMRI response in the left occipito-temporal cortex and the corresponding N1 amplitude, after brief grapheme-phoneme correspondence training in non-reading kindergarteners, improves the prediction of reading outcome over behavioral measures alone in second grade (Bach et al., 2013). No data so far exist to confirm whether the N1 ERP would also be suited for predicting reading skills before children learn letter–speech sound correspondences.

In the present article, we primarily aimed to clarify whether i) differences in N1 print sensitivity can be found in children with regular or poor reading outcomes before they learn the basic principles of reading, and ii) a pre-literate difference in the N1 print sensitivity could be used as a predictor of reading outcome at school age. We therefore report the development of this print sensitivity in a longitudinal training study from kindergarten to second grade. In contrast to our previously published articles of the same study (Brem et al., 2010; Bach et al., 2013) we i) compare the developmental trajectories of children with normal reading outcomes and children with poor reading outcomes until second grade, ii) use ERP data of an implicit word-processing task for prediction analysis, and iii) additionally compare our longitudinal training cohort with second graders who have not participated in the preschool grapheme-phoneme training in second grade. Based on previous literature, we hypothesized that the N1 print sensitivity in second graders with poor reading skills is diminished as compared to normal-reading second graders (Maurer et al., 2007). We expected that the attenuation of print sensitivity should be especially pronounced in those children who did not participate in the preschool letter-speech sound correspondence training. The search for potential predictors of reading outcome in the visual ERPs of preliterate children could finally complement and improve prediction with behavioral measures and help to identify struggling children before problems emerge.

## **Methods**

### **Study Design and Subjects**

At kindergarten age, 31 healthy, right-handed, native (Swiss-) German-speaking children (age range 5.7 to 6.9 years, mean age 6.37 ( $SD=0.3$ ), 16 girls) took part in a preschool grapheme-phoneme (GG) training session and in a longitudinal follow-up test in second grade. Note, the children, whose data are summarized in the present article, belong to the same study group as those children in our previous articles (Brem et al., 2010; Bach et al., 2013, Bach et al., 2010). The longitudinal study design included an initial comprehensive behavioral assessment, as well as EEG or EEG and fMRI recordings. The imaging sessions



took part before (“baseline”), in between, and after two preschool training periods (approximately eight weeks per training game), in which the children learned either letters and letter-sounds with computerized, non-commercial grapheme-phoneme training (“Graphogame”: GG (Lyytinen, Erskine, Kujala, Ojanen, & Richardson, 2009; Lyytinen et al., 2007; Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2011), see detailed description below) or numbers and calculations with non-linguistic control training (NC (Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009); for detailed EEG and fMRI results of the preschool training data, we refer to (Brem et al., 2010)). Both computerized training games were developed at the University of Jyväskylä and have been adjusted for the German language. All children returned for a follow-up assessment in their second school year (mean age at follow-up: 8.37 years,  $SD=0.3$ ); the assessment included behavioral tests and EEG, and some children also participated in fMRI recordings. In the present article, we focus on the following three test times: baseline (T0) before any training in kindergarten, post GG-training (T1) in kindergarten (corresponds for all children to the test time directly after GG-training: note that 15 children started with the GG intervention while 16 children started with the non-linguistic control training prior to GG-training), and second grade (T2). The children all reported normal or corrected-to-normal vision, and had a nonverbal IQ of  $>85$  estimated with the block design test of the HAWIK-III (Tewes, Rossmann, & Schallberger, 1999) in second grade. The parents of the children completed a questionnaire regarding their own reading histories (adult reading history questionnaire, ARHQ (Lefly & Pennington, 2000)) to assess the children’s familial risk of dyslexia in kindergarten. For three children, only the ARHQ of the mother was available and used for familial risk estimation. For all other children, the mean ARHQ score of both parents was computed.

Based on their word reading performance in second grade, the children participating in the longitudinal GG training study were retrospectively grouped into age-appropriate, regular readers (RR) (GG-RR:  $n=13$ , kindergarten mean age 6.43 years ( $SD=0.3$ ), second grade mean age: 8.4 years ( $SD=0.3$ ), 7 girls) or poor readers (PR) (GG-PR:  $n=12$ , kindergarten mean age: 6.35 years ( $SD=0.3$ ), second-grade age: 8.4 years ( $SD=0.3$ ), 6 girls).

A child was defined as a poor reader when he or she scored below the 25<sup>th</sup> percentile in a composite score of a standardized reading test, which included speeded word and text reading (SLRT (Salzburger Lese- und Rechtschreibtest SLRT) (Landerl, Wimmer, & Moser, 1997)) and as a regular reader when s/he scored above the 40<sup>th</sup> percentile, in accordance with our previously published articles (Bach et al., 2010; Bach et al., 2013). Children scoring between the 25<sup>th</sup> and the 40<sup>th</sup> percentiles were not included in statistical group comparisons but only in the regression/prediction analyses ( $n=6$ , 3 girls). Another standardized reading test (SLRT-II (Moll & Landerl, 2010)) was conducted with the same children in second grade. As expected, the performance in the two word-reading tests (SLRT and SLRT-II) was highly correlated ( $r=0.93$ ,  $p<0.001$ ).

The second graders of the longitudinal GG training cohort were additionally compared to a control cohort (Ctrl), i.e., matched groups (Ctrl-RR:  $n=10$ , mean age 8.25 years ( $SD=0.5$ ), 6 girls, Ctrl-PR:  $n=11$ , mean age 8.25 years ( $SD=0.3$ ), 5 girls) of second graders who had not participated in the longitudinal grapheme-phoneme training in kindergarten. This control cohort was only tested once, with a behavioral assessment and an EEG recording session in second grade; all children of the control cohort conducted one and most of the children of the longitudinal GG cohort ( $n=29$  out of 31) also participated in one or more fMRI recordings.

The regular and poor readers from the control (Ctrl-PR, Ctrl-RR) and the longitudinal (GG-PR, GG-RR) cohorts were matched for their estimated non-verbal IQ measured in second grade (block design test of the HAWIK-III (Tewes et al., 1999)). For both—the longitudinal GG and the control cohorts—the same criteria for grouping were applied.

### **Grapheme-Phoneme Training Intervention: “Graphogame”**

The intervention took place between T0 and T1 and included on average a total of about 3.6 hours of grapheme-phoneme training (plus training with a non-linguistic control game for a similar amount of time for 16 out of 31 children). The specific grapheme-phoneme training was implemented using the computerized Graphogame developed at the University of Jyväskylä (Lyytinen et al., 2009; Lyytinen et al., 2007; Saine et al., 2011). In this game, the

subject's task is to recognize and select using the mouse the visual counterpart of an auditorily presented target phoneme among several visually presented graphemes. The content and number of the choices are adapted according to the subject's performance, so as to make the training as efficient as possible. An important general principle in the design of the game is introducing the associations starting from the phonemes that are represented most consistently and frequently (e.g. /d/, /r/, /l/) by a single grapheme in a given language, advancing to less consistently represented phonemes (e.g. /t/, /p/, /f/) and phonemes represented by complex graphemes in German (e.g. /x/, /S/, /N/). Starting with consistent items avoids confusing the children with multiple incompatible orthographic representations of the phoneme, which should lead to more efficient training. Furthermore, as the connections to be learned first may require a large number of repetitions, it is important to try to avoid items that may condition connections that are incorrect in some contexts (Lyytinen et al., 2007).

In the present study, the consistency levels of the game items used were objectively determined using a novel quantitative definition of the level of consistency between phonemes and any orthographic unit. The input dataset for the analysis was formed by extracting all monosyllabic words from the German CELEX database (Baayen, Piepenbrock, & van Rijn, 1993), and the consistency level was computed for each introduced phoneme. The consistency range of the initial speech sound-letter connections to be learned was 93-100% in the first game levels, whereas consistency was somewhat lower in later levels (32-99%). When designing the levels for the game also visual and auditory similarity of introduced graphemes and phonemes was taken into account. Highly similar graphemes (e. g. d - b) or phonemes (e. g. /m/ - /n/) were thus not introduced in the same levels to avoid confusion. After introduction of phonemes with consistently or less consistently represented graphemes, the game proceeded to longer items such as consonant-vowel (CV) or vowel-consonant (VC) clusters, followed by syllables, monosyllabic words and pseudowords and finally exception words (cf. Brem et al., 2010 for a detailed description of game levels).

## **Behavioral Assessment**

**Kindergarten (baseline).** All kindergarten children participated in an initial behavioral assessment (cf. Brem et al., 2010) at their homes prior to starting the training study. We assessed reading skills (all children were non-readers before starting the training: no child could “read”/“decode” more than five high-frequency nouns out of a list of 30 given words derived from the SLRT (Landerl et al., 1997) (cf. Brem et al., 2010)), and we also evaluated precursors of later reading and spelling skills, such as rapid naming of objects (RAN), IQ (CPM: Colored Progressive Matrices (Raven, 2002)), receptive vocabulary and word comprehension (two subtests of Marburger Sprachverständnistest für Kinder, MSVT (Elben & Lohaus, 2000)), and letter knowledge (upper case “Letters UC” and lower case “Letters LC”). Further, we administered a literacy screening battery (BISC: “Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten” (Jansen, Mannhaupt, Marx, & Skowronek, 1999)) that tests phonological awareness, as well as memory and attention skills. The BISC includes the following subtests: pseudoword repetition (“BISC Pseudoword-Rep”), rhyming (“BISC Rhyme”), visual word comparison time and accuracy, phoneme association, rapid naming of congruent and incongruent colors of objects (“BISC RAN-Colours-con”, “BISC RAN-Colours-inc”), naming of colors, syllabic segmentation, and phoneme extraction. A “BISC total risk score” can be computed by the performance of a child in all subtests and indicates whether a child has a high risk for developing reading and/or spelling difficulties. Further, specific subscores of the BISC can be combined into measures reflecting phonological awareness in a broad sense (“PABs”: speech skills associated with rhyming and clapping games, defined by the summed raw scores of the subtests rhyming and syllable segmentation) or phonological awareness in a narrow sense (“PANs”: analysis of the phoneme structure without a rhythmic, segmental language context, defined by the sum of the raw scores in the subtests phoneme association and phoneme extraction (Jansen et al., 1999). The “reading/decoding” skills of the children, as well as their letter knowledge, were assessed after each intervention as well as in second grade

**Second grade.** The behavioral assessment in second grade was conducted at the children's homes. It included word, pseudoword, and text reading of the reading tests SLRT (Landerl et al., 1997) and SLRT-II (Moll & Landerl, 2010), as well as spelling (SLRT) (Landerl et al., 1997). The children's verbal and nonverbal IQs were estimated with the HAWIK-III (Hamburg-Wechsler-Intelligenztest für Kinder (Tewes et al., 1999), subtests block design (nonverbal) and similarities (verbal)). Further, we examined rapid naming of objects (RAN), letter knowledge (LK-UC and LK-LC), and phonological skills with the "BAKO" test battery ("Basiskompetenzen für Lese-Rechtschreibleistungen BAKO 1-4" (Stock, Marx, & Schneider, 2003), all subtests: pseudoword segmentation, vowel substitution, rest word identification, phoneme inversion, sound categorization, vowel length determination, and word inversion); we also evaluated the children's working memory (adapted digit span forward ("WM-forward") and backward ("WM-backward")) by having them repeat color names by means of a visual template.

The behavioral measures in kindergarten and second grade are summarized in Tables 1 and 2. Differences between regular and poor readers in those measures have been computed with independent t-tests for the longitudinal GG-training and the control cohorts separately. For the second-grade measures, an overall analysis comparing all poor versus all regular readers is given (Table 2) as well.

### **Event-related Potentials**

**Task.** The same task, testing print sensitivity (Figure 1), was conducted in all EEG (and fMRI) sessions: at the baseline assessment (T0) and after (T1) the grapheme-phoneme training in kindergarten (and after the non-linguistic control training), as well as in second grade (T2). The task tested implicit visual, auditory, and audiovisual word (W) and false font (FF)/rotated speech processing. A simple modality judgment task (press either left or right button for auditory/visual stimuli and press both buttons for audiovisual stimuli). The overall performance in kindergarten and second grade was good (accuracy (Mean  $\pm$  SD) second grade (PR and RR of GG and Ctrl cohorts) T2: 88.8  $\pm$  9.0%, FF: 90.2  $\pm$  10.3%; accuracy kindergarten (all children) T0 W: 74.9  $\pm$  19.0%, FF: 72.5  $\pm$  20.3%, T1 W: = 75.6  $\pm$  17.5%, FF:

70.5± 18.2%) even though some (n=12 out of 31) kindergarten children performed on chance level or below in one or both conditions at T0 and/or T1. Despite poor performance, all children were included because the modality judgment task was used primarily to keep the children focused on the stimulus. Words and false font stimuli were presented for 850ms in black on a white background in the center of the screen. During the inter-stimulus-interval (ISI) (2650ms) and during the null events (omitted stimuli or no-stimulus events), a centered fixation cross appeared. During the entire task, icons of an eye and an ear were shown on the left and right sides of the screen, respectively, in order to remind children which button to press. Mean stimulus onset asynchrony (SOA) added up to 3500ms for both real stimuli and null events. The ERP task included the pseudorandomized presentation of 42 stimuli per condition and a total of 42 null events. Word and false font stimuli were matched for the main visual characteristics, such as character size, string length (3-5 characters), and number of ascenders and descenders in a string. The stimuli were either presented unimodally (visual or auditory) or bimodally (visual and auditory simultaneously), whereby the visual and auditory stimuli could be congruent (same words) or incongruent (different words). The analyses of the present article focus on the ERPs to purely visual conditions (W, FF) at baseline (T0), post GG-training (T1), and the follow-up in second grade (T2). Note that we do not include the data assessed after the control training in kindergarten (for results please consult Brem et al., 2010).

## **Behavioral Statistics**

**Development and training.** To examine the effects of training in kindergarten and development on accuracy and reaction time (RT) in our implicit task, we computed multivariate analyses of variance (MANOVAs) for the longitudinal GG training cohort with between-subject factor of reading level (RR/ PR) and within-subject factor of test time (T0, T1, T2) and condition (W, FF) for accuracy and reaction time, separately.

**Regular vs. poor readers in second grade.** To compare performance between RR and PR, we computed separate MANOVAs for RT and accuracy, with between-subject

factors of cohort (GG/ Ctrl), reading level (RR/ PR), and the within-subject factor condition (W, FF).

### **Recording and Processing of EEG Data**

The ERPs of all children and at all three test times (T0, T1, T2) were recorded from 64 channels with filters set at 0.1-70Hz and a sampling rate of 500Hz. Impedances were kept below 15kOhm. The EEG montage included all 10-20 system electrodes, plus the following additional electrodes: FPz, FCz, CPz, POz, Oz, Iz, AF1/2, F5/6, FC1/2, FC3/4, FC5/6, FT7/8, FT9/10, C1/2, C5/6, CP1/2, CP3/4, CP5/6, P5/6, TP7/8, TP9/10, PO1/2, PO9/10, OI1/2, PPO9h/10h, and two EOG electrodes below the outer cantus of each eye. For more even coverage, the four electrodes O1', O2', Fp1' and Fp2' were placed 15% more laterally to Oz or Fpz, respectively. The electrodes OI1 and OI2 were placed to the left and to the right of the midline, halfway between Oz and Iz, to provide better coverage of the occipital scalp distributions.

Children were seated in front of a computer screen wearing headphones for auditory stimulation. The continuous EEG data were first down-sampled to 256Hz. To identify and correct ocular artifacts caused by eye blinks, or by lateral or horizontal eye movements, an independent component analysis (ICA (Jung et al., 2000)) on filtered (0.1-30Hz) data was computed and the corresponding components removed. After correcting for ocular artifacts, the data were bandpass filtered (0.3-30Hz, 24dB), epoched (-125 to 1125ms), and re-referenced to average reference (Lehmann, 1984). Before computing condition averages, all epochs with artifacts exceeding  $\pm 100\mu\text{V}$  (for two kindergarten children with high raw EEG, an artifact criterion of  $\pm 125\mu\text{V}$  was used) at any channel were rejected. Separate group averages were computed for poor- and normal-reading children of the longitudinal training and the control cohorts and separately for kindergarten (baseline and post GG training) and second-grade children. All condition averages for each child included at least 15 epochs each (Mean $\pm$ SD: Kindergarten, T0, W:  $36.3\pm 5.9$ ; FF:  $34.9\pm 5.9$ ; T1: W:  $35.3\pm 5.3$ ; FF:  $35.8\pm 5.1$ ; Second grade, T2 (GG), W:  $37.6\pm 6.5$ ; FF:  $37.1\pm 5$ ; and T2 (Ctrl), W:  $36.8\pm 4.4$ , FF:  $38.7\pm 2.4$ ). Note, there was no significant main effect or interaction regarding the number of segments

included in the averages, according to the multivariate analysis of variance (MANOVA) with factors condition (W, FF), test time (T0, T1, T2), and reading level (PR/RR).

**N1 amplitude analyses.** The N1 interval was determined by means of two subsequent global field power (GFP) sinks (188–289 ms) in the grand average waveform, computed as the mean of the averages to W and FF for all groups and all analyzed test times (kindergarten baseline and post-training, and second grade). The mean amplitude value within this interval has been determined for a left “LOT” (comprising the electrodes: PO9, PPO9h, O1', P7) and for the corresponding right “ROT” (PO10, PPO10h, O2', P8) occipito-temporal electrode cluster, similar to our previous article (Bach et al., 2013).

## **ERP Statistics**

**Development and training.** To examine the effects of the training and development on N1 amplitude and lateralization, we computed a repeated measures MANOVA for the longitudinal cohort with between-subject factor reading level (RR/ PR) and within-subject factors time (kindergarten baseline (T0), kindergarten post-training (T1), second grade (T2)), condition (W, FF), and hemisphere (left, right).

**Regular versus poor readers in second grade:** To compare the two cohorts of second graders, we used a MANOVA with between-subject factors of cohort (GG/Ctrl) and reading level (RR/ PR), and the within-subject factors of condition and hemisphere.

**Prediction analysis.** For the whole longitudinal training cohort ( $n = 31$ , including the six children with intermediate reading scores), we also conducted a stepwise multiple regression analysis (adding  $p < 0.05$ , keeping  $p < 0.1$  significant predictors) to predict second-grade reading skills with behavioral and ERP measures collected in kindergarten (T0). Note that in our previously published article (Bach et al., 2013), we used ERP and fMRI measures of an explicit word processing task collected after Graphogame training (T1) for prediction. As the dependent (outcome) variable, we always used speeded word reading determined by the SLRT-II in second grade (Table 3). All analyses were also repeated for the outcome measure of pseudoword reading (SLRT-II) and are detailed below:



**Analysis 1.** In a first step, we entered all behavioral variables (except for composite measures such as PABs, PAns, and BISC score) from baseline kindergarten assessment as well as age, IQ, ARHQ, and parents' estimated SES to examine the behavioral variables that best predict word reading in second grade.

**Analysis 2.** As a next step, we determined the predictive power of neurophysiological measures for the same outcome variables. Our approach was based on the finding that the print-sensitive N1 negativity over left hemispheric occipito-temporal channels, together with the left fusiform activation after Graphogame training in an explicit word processing task, contributed to prediction of the reading score in the same children (Bach et al., 2013). Here, we tested whether the N1 negativity over the left and right hemispheres, evoked in an implicit word- and falsefont-processing task before grapheme-phoneme training, would also predict second-grade reading skills. In addition to the print sensitive negativity (W-FF) over LOT and ROT, we tested the predictive value of the N1 mean amplitude to words and falsefonts before Graphogame training.

**Analysis 3.** Finally, we combined behavioral and neurophysiological predictors in a third multiple regression analysis. Predictor variables that significantly contributed to predictions in Analysis 1 or 2 were entered as separate blocks to examine whether the combination yields an increase of the explained variance in reading measures.

## **Results**

We report statistically significant results ( $p < 0.05$ ), trends of specific interest ( $p < 0.1$ ) and also include partial eta squared ( $\eta_p^2$ ) as a measure of effect size:  $\eta_p^2$  reflects the proportion of variance (effect plus error variance) explained by a specific effect or interaction which is not explained by other factors in the analysis.

### **Behavioral Data**

**Behavioral assessment in kindergarten and second grade.** The results of the behavioral test batteries conducted in kindergarten before the start of the grapheme-

phoneme training (T0), as well as the data of the four groups of second graders (T2), are listed in Tables 1 and 2.

**Kindergarten.** Future poor- and normal-reading kindergarteners did not differ regarding age, gender, and nonverbal IQ. Future poor-reading children performed worse in reading-related skills such as lowercase letter knowledge, receptive vocabulary, phonological awareness (PABs, rhyming, phoneme association), and rapid automatized naming.

**Second grade.** In addition to all reading and spelling measures, poor and regular readers of both the longitudinal GG and control cohorts also differed in their phonological skills (as indicated by the BAKO test screening battery), the verbal IQ (estimated with the HAWIK-III subtest similarities), lowercase letter knowledge, and RAN (trend for Ctrl cohort). Note that in the separate univariate ANOVAs with between-subject factors of reading level (RR, PR) and cohort (GG, Ctrl) for the core behavioral variables in second grade (three measures for reading, phonology (BAKO score), RAN, verbal IQ, spelling, letter naming), no significant interaction was found between cohort and reading level. Reading level was significant for all behavioral variables. A main effect for cohort pointed to better spelling skills in the longitudinal sample ( $F(1,42)=5.23$ ,  $p=0.027$ ), but the significance level did not exceed the Bonferroni corrected threshold of  $p<0.0063$ .

### **ERP Task Performance**

**Development and training.** Children improved their response accuracy in the modality judgment task with test time ( $F(2,22)=19.46$ ,  $p<0.001$ ,  $\eta_p^2 = 0.64$ ) and also became faster ( $F(2,22)=3.92$ ,  $p=0.035$ ,  $\eta_p^2 = 0.26$ ). No other main effect or interaction reached statistical significance.

**Regular versus poor readers in second grade.** No significant condition or cohort differences in accuracy and reaction times were found in second graders.

### **ERP Data**

**Development and Training (Figures 2 & 3).** The repeated measure MANOVA on N1 amplitude with between-subject factor of reading level (regular, poor) and within-subject factors of test time (T0, T1, T2), condition (W, FF), and hemisphere (left, right) revealed

significant main effects of test time ( $F(2,22)=7.1$ ,  $p=0.004$ ,  $\eta_p^2=0.39$ ) and condition ( $F(1,23)=56.5$ ,  $p<0.001$ ,  $\eta_p^2=0.71$ ), as well as interactions of test time x condition ( $F(2,22)=16.7$ ,  $p<0.001$ ,  $\eta_p^2=0.60$ ) and test time x reading level ( $F(2,22)=5.8$ ,  $p=0.010$ ,  $\eta_p^2=0.34$ ). Additional two trends for test time x hemisphere x reading level ( $F(2,22)=2.68$ ,  $p=0.091$ ,  $\eta_p^2=0.2$ ) and test time x hemisphere x condition x reading level ( $F(2,22)=2.8$ ,  $p=0.085$ ,  $\eta_p^2=0.2$ ) indicated that the N1 print sensitivity development tended to differ between regular readers and poor readers and depended on hemispheres.

To explain the fourfold interaction in more detail, we performed post-hoc MANOVAs for all three test times separately. None of these post-hoc MANOVAs pointed to a significant main effect of reading level or interaction between reading level and condition. The main effect of condition (W>FF) remained significant at all test times. Only in kindergarteners (T0) did the triple interaction of reading level x hemisphere x condition show a trend ( $F(1,23)=3.2$ ,  $p=0.088$ ,  $\eta_p^2=0.12$ ), whereby future poor readers displayed a more pronounced negativity to words over the right rather than the left hemisphere (post-hoc t-tests: words:  $t=2.67$ ,  $p=0.022$ , trend for print sensitivity:  $t=2.04$ ,  $p=0.066$ ), while regular readers showed similar amplitudes for both conditions and the condition difference over both hemispheres (all t-tests,  $p=ns$ ). The corresponding statistical topographic maps (Figure 3A) nicely illustrate this finding; no significant print sensitivity over occipito-temporal channels was found in future regular readers, but a pronounced right lateralized print sensitivity was detected in the group of future poor readers before training.

**Regular versus poor readers in second grade (Figures 3B & C).** The MANOVA with between-subject factors of cohort (GG vs. Ctrl) and reading level (RR vs. PR) and the within-subject factors of condition (W, FF) and hemisphere (left, right) revealed an interaction for the two between-subject factors, cohort x reading level ( $F(1,42)=5.3$ ,  $p=0.027$ ,  $\eta_p^2=0.11$ ). The interaction reflected a more pronounced difference in the overall N1 amplitude of the poor reading GG and Ctrl cohorts, as compared to the difference of the two cohorts of regular readers (Figure 3C).

Further, a significant main effect of condition ( $F(1,42)=101.5$ ,  $p<0.001$ ,  $\eta_p^2=0.71$ ) and an interaction of condition with hemisphere ( $F(1,42)=4.3$ ,  $p=0.044$ ,  $\eta_p^2=0.093$ ) was found. The N1 amplitude was more pronounced for words than falsefonts over left than right occipito-temporal sites. Given the significant interaction between reading level and cohort (GG, Ctrl), we conducted post-hoc MANOVAs and analyzed the N1 amplitudes for the GG and the Ctrl cohorts separately. For these MANOVAs, reading level was used as a between-subject factor. The GG ( $F(1,23)=63.3$ ,  $p<0.001$ ,  $\eta_p^2=0.73$ ) and the Ctrl ( $F(1,19)=42.7$ ,  $p<0.001$ ,  $\eta_p^2=0.69$ ) cohorts showed a main effect for condition; the Ctrl cohort also showed a trend for hemisphere  $\times$  condition ( $F(1,19)=3.9$ ,  $p=0.062$ ,  $\eta_p^2=0.23$ ). In addition, our topographic statistical maps (Figure 3B) point to a significant difference between the print sensitivity of regular and poor readers in the control cohort at left parieto-occipital sites. No difference in the N1 print sensitivity distribution between the reading level groups was found for the GG training cohort.

### **Prediction Analyses with Behavioral and Neurophysiological Measures**

**Analysis 1.** In a first step, we determined the behavioral parameters in kindergarten (derived from the initial assessment) that best predict childrens' reading outcomes in second grade. As the dependent (outcome) variable, we always first looked at word reading (number of correctly read words per minute: SLRT-II). We then repeated the analysis for the outcome variable "pseudoword reading" (SLRT-II). Predictors were entered stepwise and included age (at assessment), an estimate of the parents' SES, all subtests of the BISC, rapid object naming (RAN), the familial risk score determined by the ARHQ, two measures for vocabulary (receptive vocabulary and comprehension), letter knowledge (upper and lower case), estimated IQ, and a measure for preschool word reading skills. No multicollinearity ( $r \geq .9$ ) was found between predictors. The stepwise procedures (Table 3) showed that the two measures, RAN and estimated IQ (nonverbal), were kept for the model and significantly contributed to the prediction of second-grade word reading (combined model:  $R^2=0.32$ , adjusted  $R^2=0.27$ ,  $p=0.005$ ), while only RAN significantly explained variance in pseudoword reading ( $R^2=0.28$ , adjusted  $R^2=0.26$ ,  $p=0.002$ ).

**Analysis 2.** As a next step, we analyzed the predictive value of electrophysiological measures. The same two outcome measures for behavioral analyses (word, pseudoword reading in second grade) were entered in the stepwise multiple regression analyses. As predictors, we used the N1 mean amplitude measures to words and falsefonts and their difference (W-FF) over left or right occipito-temporal sites (LOT or ROT) at the baseline test in kindergarten. The baseline N1 print sensitivity over ROT accounted for 22.1% ( $R^2 = 0.25$  adjusted  $R^2 = 0.22$ ,  $p = 0.004$ ) of the variance in later word reading or 20% of later pseudoword reading ( $R^2 = 0.23$ , adjusted  $R^2 = 0.20$ ,  $p = 0.007$ ). Poor reading in second grade was thus best explained by the enhanced N1 negativity to W versus FF over the right hemisphere in kindergarten. The correlation between the electrophysiological predictor and second-grade word reading indicated that the less negative the W-FF difference over right occipito-temporal sites, the better were children's future reading skills in second grade (see Figures 3A and 4).

**Analysis 3.** As a last step, we analyzed whether the electrophysiological predictors improve the prediction of reading outcomes over behavioral data alone. Based on our previous analyses, we thus entered RAN and estimated IQ as a first step and the differential N1 negativity (W-FF) over the right hemisphere as a second step in the final multiple regression analysis. The combination of the behavioral and the neurophysiological measures yielded the best prediction for word reading ( $R^2 = 0.46$ , adjusted  $R^2 = 0.4$ ,  $p < 0.001$ ) and pseudoword reading ( $R^2 = 0.45$ , adjusted  $R^2 = 0.39$ ,  $p < 0.001$ ), and together, they explained 35% to 40% of the respective variance. The significant change in the  $R^2$  (word reading:  $\Delta R^2 = .15$ ,  $p = 0.012$ ; pseudoword reading:  $\Delta R^2 = .14$ ,  $p = 0.014$ ) demonstrated that prediction improved by adding the electrophysiological measure.

## **Discussion**

This is the first study to investigate the predictive value of early visual ERPs in pre-literate preschool children for prognosis of reading outcome. A considerable number of studies have already identified behavioral predictors of language and reading outcomes at preschool age and pointed to a number of measures that may be used to estimate a child's

risk for potential reading problems. Previous research has also comprehensively studied the development of print sensitivity in the brain from kindergarten to adulthood in children with normal and poor reading achievement or during a computerized grapheme-phoneme training in non-reading kindergarteners (Brem et al., 2010). In our latest article, we demonstrated that the N1 print sensitivity initiated by preschool grapheme-phoneme correspondence training—along with the corresponding VWFS activation—complements and improves the prediction of reading outcome in second grade over behavioral data alone (Bach et al., 2013). Data about the predictive value of the visual N1 to print at preschool age before any reading related training, however, are still lacking, even though the high predictive value of auditory ERPs have long been known (Guttorm et al., 2005; Guttorm et al., 2010, Leppanen et al., 2010, Maurer et al. 2009, Molfese et al 2000). Similar to auditory ERPs, visual ERP measures such as the N1 are easy and relatively cheap to be collected in kindergarten children and might be especially useful for complementing behavioral or ERP test batteries for early identification of children with poor reading outcomes.

Here, we first report the changes in the N1 amplitude and topography for two groups of children with normal and poor reading achievements before and after an eight-week computerized letter-speech sound training in kindergarten, as well as at the longitudinal follow-up in second grade. Additionally, we compare the second-grade data of our longitudinal GG training groups with age-, IQ-, and reading-matched peers who have not participated in the grapheme-phoneme preschool training. Most importantly, we examine whether an early, pre-literate print-sensitive ERP, assessed prior to learning to read, may contribute to predictions of reading outcomes at school. In contrast to our previous article on the prediction of reading outcome by means of the print-sensitive responses measured with ERP and fMRI in an explicit word-decoding task (Bach et al., 2013), we analyzed the data of an implicit audiovisual print and falsefont/rotated speech-processing task. Further, we concentrated on predictive measures at the baseline test in kindergarten before any grapheme-phoneme correspondence or other reading-related training.

### **Initialization and Establishment of Print Sensitivity when Reading Skills Are Acquired**

As already reported in our previous article (Brem et al., 2010), the data largely confirm the findings of past studies regarding a more pronounced N1 to words than false fonts developing when children learn to read (Maurer et al., 2005; Maurer et al., 2006). The more pronounced left occipito-temporal negativity to words than false fonts or symbols strings is absent in non-reading kindergarten children prior to learning grapheme-phoneme correspondences (Brem et al., 2010). The development of the typical left occipito-temporal N1 print sensitivity is most likely initiated by processes linking graphemes and phonemes. Accordingly, it becomes only apparent when children start to learn associations between letters and speech sounds (Maurer & McCandliss, 2007). Data of a computerized grapheme-phoneme training for non-reading kindergarteners (Brem et al., 2010) and data of an artificial script training for adults (Yoncheva, Blau, Maurer, & McCandliss, 2010) provided evidence of this hypothesis. Both studies suggested that grapheme-phoneme learning and not visual familiarity alone initiates the characteristic left hemispheric print-sensitive negativity seen in the N1. One has to bear in mind that it is difficult to disentangle and weight the impact of visual familiarity and grapheme-phoneme learning on the development of the N1 print sensitive response. Given that the characters (letters, symbols or character strings) are repeatedly presented, the training procedures automatically involve a familiarization process with the visual appearance of the novel script.

The general principle of the training game used for the present study was to introduce the phonemes that are most consistently and frequently represented by graphemes at the beginning, whereas phonemes with more inconsistent connections to orthographic symbols were presented later. As the children's prior exposure to more inconsistent grapheme-phoneme associations cannot be completely avoided, the principle of starting the training with the most consistent associations—as well as its adaptive design—may have had an effect on the efficiency of the training and could also explain the corresponding rapid learning effects demonstrated in the ERP data in poor and normal readers. More direct comparisons, however, would be needed to allow judging about the efficiency of the Graphogame in

comparison to other preschool trainings and to conclusively value the importance of the consistency and frequency principle in learning to read alphabetic languages.

Interestingly, a pronounced pre-literate and atypical print sensitive negativity over the right hemisphere was detected in future poor-reading children at baseline. This atypical right hemispheric distribution of the pre-literate negativity tended to be stronger in future poor than future regular readers. Potential explanations and implications of this finding are discussed in the next section. The minor topographical differences in the N1 print sensitivity distribution between reading level groups at baseline disappeared with grapheme-phoneme training. After the training, both groups showed a clear differentiation between conditions and similar N1 distributions with the typically more pronounced negativity to print over the left occipito-temporal cortex. Against our expectations, no significant difference in the N1 print sensitivity over left or right occipito-temporal channels between regular and poor readers was found in the second-grade data of our GG training cohort. This clearly contrasts with previous studies in which an attenuation of print sensitivity in second graders, pre-adolescent children or adults with dyslexia was found (Araujo et al., 2012; Maurer et al., 2007, Mahé et al., 2012). We cannot exclude that we missed this N1 attenuation because we compared normal-reading children to poor readers (below the 25<sup>th</sup> percentile), instead of to children with developmental dyslexia. When applying more stringent reading performance criteria (below the 10<sup>th</sup> percentile) for examining the poorest children in the present study, the number of children became too small for proper analysis. Only in a subsidiary analysis did we compare the N1 map topographies between our 13 regular readers and the seven children with very poor reading skills. However, even in this group of very poor-reading second graders, there was no indication of deficient print sensitivity in the N1 ERP. In contrast to our GG training cohort, we found a significant difference when comparing the N1 print sensitivity topographies of poor and regular readers in the control cohort, who did not participate in the preschool GG training. The comparison of the mean N1 amplitudes over LOT and ROT in the MANOVA did not yield such a difference, most likely because the topographic group



difference was located at somewhat more dorsal electrode sites than those included in LOT. The statistical maps showed a clear reduction of the print-sensitive negativity over left posterior parieto-temporo-occipital scalp sites in poor readers of the control cohort. This reduction in left hemispheric print sensitivity thus corresponds with the reduction in print sensitivity of children with dyslexia in previous studies (Araujo et al., 2012; Maurer et al., 2007). According to the dual route model of reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), grapho-phonological decoding is a key process in reading alphabetic languages. It is especially important for children, as long as they rely on the indirect path of written word processing and cannot directly access the word meaning through the word form. The lack of diminished print sensitivity in our longitudinal training group of poor readers suggests that the preschool training may have consolidated grapheme-phoneme correspondences, facilitated grapho-phonological decoding, and thereby compensated for the N1 amplitude differences. Our control cohort of second graders did not participate in the grapheme-phoneme training; the diminished print sensitivity seen in the topographic maps of the poor reading children in this cohort may thus reflect a failure in grapho-phonological decoding processes. One could also argue that phonological processes modulating the visual N1 may lose their impact when children start to access lexical information directly. Dyslexic children with more reading experience for example (fifth grade) did not differ anymore in terms of their N1 print sensitivity when compared to age-appropriate reading children in an implicit reading task (Maurer et al. 2011). It seems not plausible though, that the short GG training at preschool age accelerated the transition from indirect to direct lexical access and thus diminished the group difference in the GG cohort. A more plausible alternative explanation for the diminished (control cohort) or even absent (GG-cohort) reduction in the print sensitive N1 between poor and regular readers in the present study is the use of a different implicit reading task as compared to previous studies. The one-back repetition detection task in the study of Maurer and colleagues (Maurer et al. 2007; Maurer et al. 2011) was probably more challenging and forced the children to process the visual information in more detail than the modality judgment task used here. One would anticipate

that processing differences between groups are amplified by more challenging print-processing tasks. The lack of group differences in the N1 print sensitivity in fifth grade (Maurer et al., 2011) would coincide with this hypothesis as well because the implicit one back repetition detection task may have been less challenging for these children due to their advanced reading skills. Differences that are attributable to impaired phonological processing should appear in more demanding tasks, such as those requiring explicit processing (Mahé et al., 2012), phonological decisions, or manipulations of words and pseudowords (Kast, Elmer, Jancke, & Meyer, 2010), the more practiced the children are. Whether the difference in print sensitivity between children with and without dyslexia in the study by Araujo (Araujo et al., 2012) was caused by the somewhat more challenging implicit print-processing task (determining which of two characters was present in the previously shown stimulus) or by the fact that children with less and more reading experience have been pooled cannot be answered conclusively.

### **Behavioral Markers of Reading Outcome**

A key aim of our longitudinal study was to examine the potential of behavioral and neurophysiological measures at kindergarten age for prediction of reading outcome two years later in second grade.

All behavioral measures assessed in kindergarten—as well as age, familial risk, and an estimate of the parents' socio-economic status—were examined. From a behavioral perspective, a strong correspondence to a large number of previous studies was found: slower rapid automatized naming as well as deficits in phonological awareness, letter name knowledge, and receptive vocabulary characterized our kindergarteners with poor reading outcomes (Catts et al., 2001; de Jong & van der Leij, 1999; Lyytinen et al., 2004; Lyytinen et al., 2007; Puolakanaho et al., 2007; Scarborough, 1990; Wimmer, Mayringer, & Landerl, 2000; Wolf, 1986). Along with the highly expected deficits in kindergarten, our poor-reading second graders also exhibited the expected deficits in verbal IQ, spelling, rapid automatized naming, letter name knowledge, and several phonological tests.

From all analyzed behavioral measures in kindergarten, only rapid naming of objects and the estimated nonverbal IQ yielded significant predictions. Together, these two measures explained 27% of the variance in later reading skills. In contrast to the RAN, the estimated IQ only added a relatively small amount of explained variance (trend) when computing the combined regression analysis with behavioral and electrophysiological measures. This result is in accordance with a row of studies indicating that rapid naming is one of the most reliable predictors of reading outcome at preschool age (Wolf, 1986), especially in relatively consistent languages such as German (Wimmer et al., 2000), Finnish (Puolakanaho et al., 2007), or Dutch (de Jong & van der Leij, 1999).

### **Atypical N1 Print Sensitivity as a Neuroimaging Marker to Predict Reading Outcome**

When predicting word reading with electrophysiological measures of kindergarteners at baseline, only one measure—namely, the mean N1 amplitude of the print-sensitive right occipito-temporal negativity—significantly (22%) explained variance in later reading skills. An increase in the N1 amplitude has been attributed to gaining visual expertise to animals, objects (Gauthier, Curran, Curby, & Collins, 2003; Scott, Tanaka, Sheinberg, & Curran, 2006; Tanaka & Curran, 2001), or symbol strings (Brem et al., 2005). A more pronounced N1 negativity over the right hemisphere has usually been related to the processing of non-linguistic stimuli and, specifically, faces (Allison, Puce, Spencer, & McCarthy, 1999; Bentin, Allison, Puce, Perez, & McCarthy, 1996; Rossion, Joyce, Cottrell, & Tarr, 2003). However, pronounced right occipito-temporal negativities also occurred for linguistic stimuli. Adults showed an increase in the right occipito-temporal N1 after artificial script training (Maurer et al., 2010). Furthermore a more pronounced negativity over the right hemisphere to words than symbol strings was revealed in non-reading preschoolers with especially high letter knowledge (Maurer et al., 2005). This right occipito-temporal negativity has therefore been suggested to reflect a neural correlate of visual familiarity with letters, the elements of print. Visual familiarity may develop as a result of the abundant exposure to letters and print in children's environment (Maurer et al. 2005). Assumably, it thus develops before reading instruction starts and before children are able to read printed words. In line with our data, a

similar right lateralized negativity has also been reported for kindergarten children with poor reading outcomes (Maurer et al., 2007). A supplemental analysis yielded no correlation between letter knowledge and the N1 print sensitivity over left or right hemispheres in our kindergarteners. It is therefore questionable whether visual familiarity alone can explain this atypical print-sensitive N1 distribution. This distribution could, however, reflect a transient stage in literacy acquisition and development preceding reading onset and reflecting both visual familiarity and growing expertise with letters or letter strings. Children with normal reading outcomes may show such a distribution at a younger age. Whether the right lateralized negativity develops later in children with a poor reading outcome and reflects a delayed development of visual familiarity with print is presently not known. This question has needs to be examined in future longitudinal studies, starting with younger children. Otherwise, one could also argue that pre-literate children with poor reading outcomes use a different strategy to process print even before they learn to read. Artificial script training in adults induced an increase of the left hemispheric N1 negativity when the training focused on grapheme-phoneme correspondences, but modulated the right hemispheric negativity when whole word associations were practiced (Yoncheva et al., 2010). From these results, one could infer that the focus of attention in our children with poor reading outcomes is centered more on larger (whole words) than smaller word units. The corresponding N1 distribution thus consequently resembles the ERP topography typically seen for objects, faces, or pictographs (Rossion et al., 2003; Zhang et al., 2011) more closely. A focus on whole words may impair children's ability to learn to read alphabetic languages later on, as supported by the finding of a strong positive correlation of the differential N1 amplitude over the right hemisphere at preschool age and the number of words read two years later. Such children would first have to learn that words consist of single grapheme units, which have to be translated to phonemes in order to finally access the word meanings.

### **Prediction Achieved by Combining Neuroimaging and Behavioural Measures**

Importantly, most variance in second-grade reading was explained by combining both the behavioral and electrophysiological kindergarten measures; a notable prediction of second-grade reading skills was obtained by explaining 40% of the variance. Even though this contribution of the visual N1 ERP to prediction in this study is lower as compared to the one by auditory ERPs (Guttorm et al., 2010; Guttorm et al., 2005; Maurer et al., 2009; Molfese, 2000), this result clearly shows the potential of combining measures from different methods—such as, for example, behavior and ERPs—to advance the prediction of future readers at preschool age. It remains to be clarified whether the combination of specific ERPs from the auditory and visual modality would also further explain variance and allow for early prognosis.

### **Conclusion**

In summary, the present study suggests that a short and specific grapheme-phoneme training in kindergarten initiates durable print sensitivity in future poor and regular readers reflected by more pronounced N1 amplitudes to words than falsefonts over left occipito-temporal sites.

More importantly, our data also indicate that print is processed differently in non-reading kindergarten children with regular or poor reading outcomes. The atypical and more pronounced print sensitive N1 over the right hemisphere in preliterate children even contributed to prediction of reading outcome at school age. Whether this right lateralized print sensitive response in future poor readers reflects a developmental delay of a common, transient process during the development of print processing or, instead, represents a different and probably less successful processing strategy still needs to be clarified. However, our data provide evidence that the right lateralized print sensitive negativity in children prior to learning to read serves as a risk factor for emerging reading problems and may be used as a biomarker to complement and improve predictions of reading outcomes at an early age. Our data, together with the increasing number of studies demonstrating the contribution of specific structural and functional MR measures in predating reading outcome (Hoeft et al., 2011; Hoeft et al., 2007; Raschle et al., 2010), thus support a multimodal approach. The

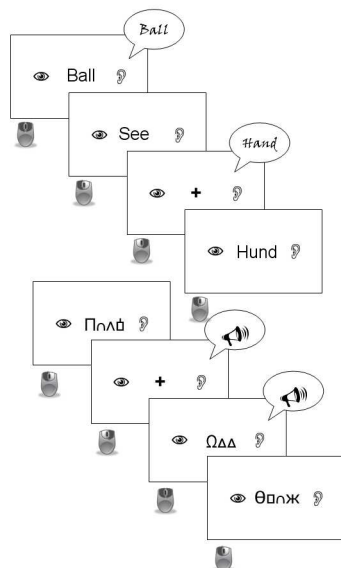
combination of measures from different methods and modalities such as behavior, ERP, and (f)MRI may optimize long-term prediction in the future.

### Acknowledgements

We thank the developing team of the “Graphogame” at the University Jyväskylä and, in particular, Jane Erskine, Anne Mönkkönen, Marika Peltonen, and Gonny Willems for implementing the training game. This work was supported by the European Commission’s FP6, Marie Curie Excellence Grants (MEXT-CT-2004-014203), the Centre of Excellence of Learning and Motivation Research, Academy of Finland, University of Jyväskylä, and the Hartmann Müller-Stiftung für medizinische Forschung (project No. 1252) of the University of Zürich.

### Figures

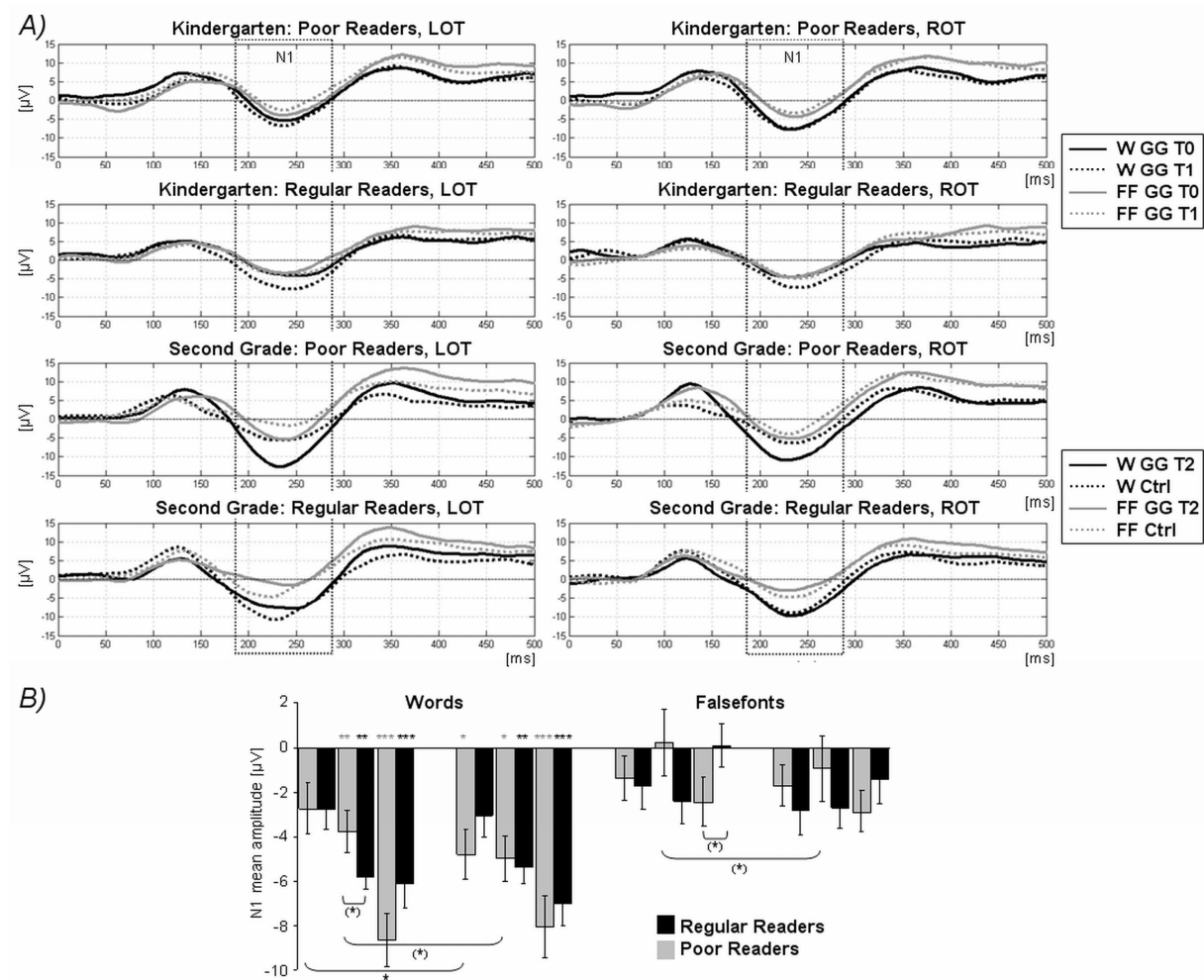
**Figure 1**



*Implicit audiovisual word and false font/rotated speech-processing task: Children had to decide by pressing a button on the modality (auditory, visual or audiovisual) of the presented*

stimulus. The experiment included two parts, in which either written and spoken words or falsefonts and rotated, non-intelligible speech stimuli were presented.

**Figure 2**

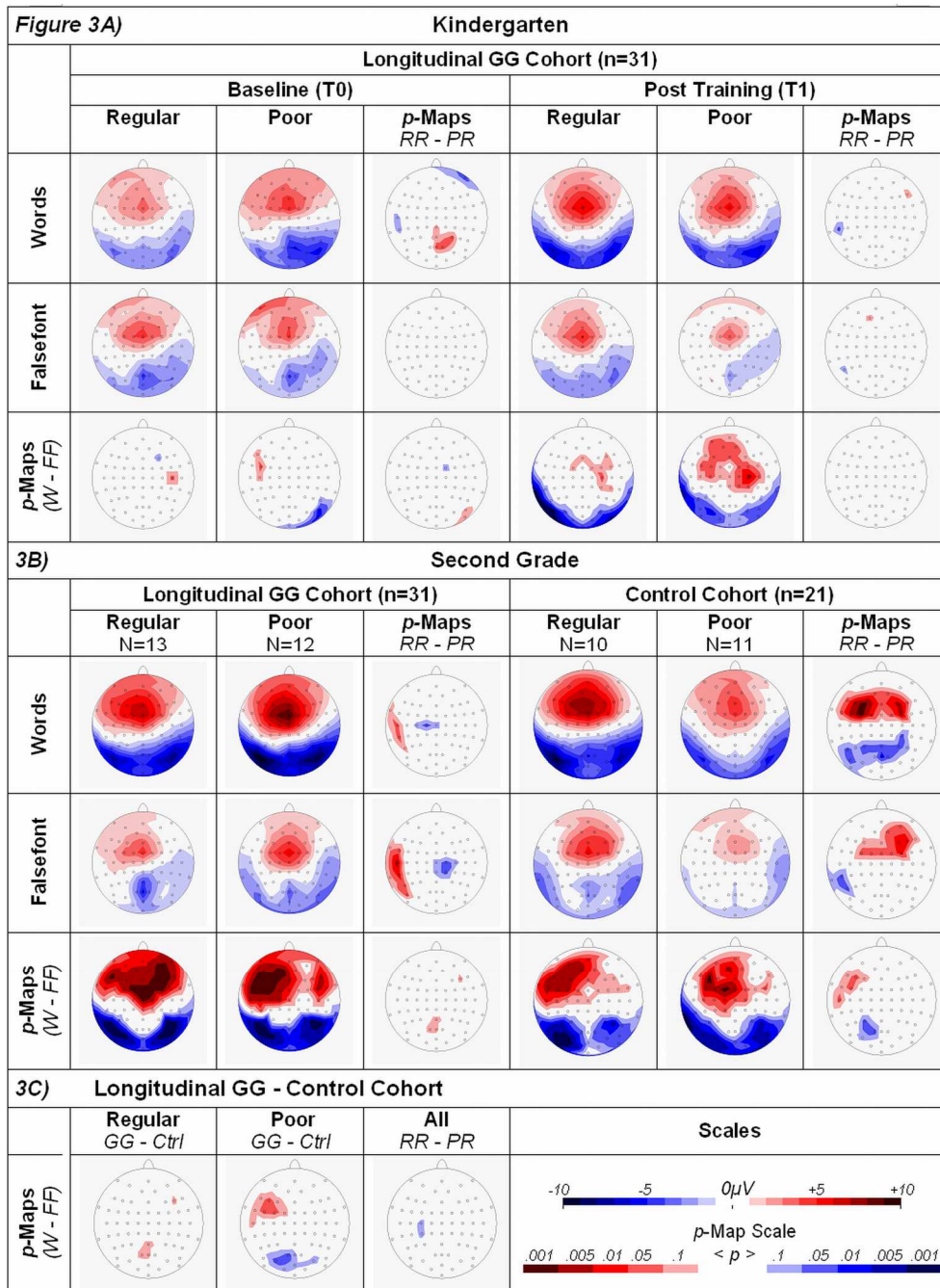


A) Waveforms at left (LOT) and right (ROT) occipito-temporal sites shown for regular readers and poor readers in kindergarten and second grade. The dotted box frames the analyzed N1 (188-289ms) interval. Words (W, black) and falsefonts (FF, grey) are shown for all test times (T0, T1, T2) and for the longitudinal training (GG) and control (Ctrl) cohorts.

B) Bar chart illustrating the mean N1 amplitudes (188-289ms) to words (left) and falsefonts (right) over left (LOT) and right (ROT) occipito-temporal electrodes for regular (black: RR) and poor (grey: PR) readers of the longitudinal GG-training cohort at each test

time (T0, T1, T2). Note, the significance level of the condition difference (W-FF) is indicated above the bars to W, in black for RR and grey for PR. \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$ , (\*)  $p < 0.1$ . Error Bars represent  $\pm 1$  SEM.

**Figure 3**



A) Potential field maps (topographies) and statistical maps (p-maps) of the longitudinal GG training cohort at baseline and post-training sessions in kindergarten. The potential field maps of the visual N1 (188-289ms) to words (top), falsefonts (middle) and the

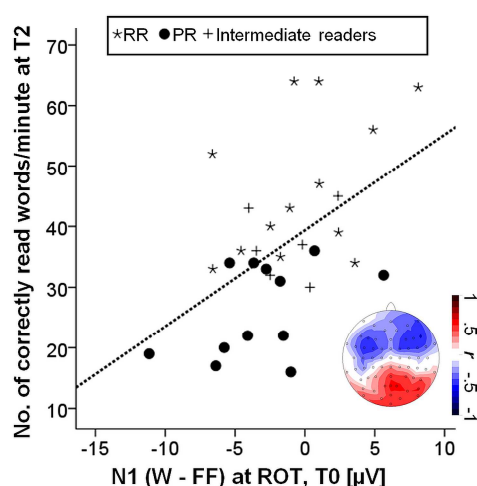


statistical condition (W-FF) differences (bottom) are illustrated for future regular readers (left), future poor readers (middle) as well as for the group difference (right). The bottom line shows the statistical condition differences and nicely displays the development of print sensitivity with grapheme-phoneme training and the right lateralized and atypical pre-literate print sensitive negativity in future poor readers at baseline.

B) Potential field maps (topographies) and statistical maps (p-maps) of the longitudinal GG-training cohort and the control cohort of age-appropriate and poor-reading second graders. The topographies are shown for each condition separately as well as for the statistical condition difference (W-FF). Both, regular and poor readers exhibited a pronounced difference between words and falsefonts in second grade. The print sensitivity was attenuated only in poor readers of the control cohort.

C) The left and middle p-maps illustrate the direct statistical comparison of the N1 print sensitivity between the longitudinal GG-training and control cohorts, for regular and poor readers, respectively. The left occipito-temporal print sensitive negativity is more pronounced in poor readers of the longitudinal GG-cohort than the control cohort. On the right, the statistical difference in print sensitivity between all regular and all poor readers (pooled over longitudinal GG-training and control cohorts) is illustrated.

**Figure 4**



Correlation of the mean N1 amplitude (W-FF) over right occipito-temporal sites in kindergarten (baseline) with second grade word reading speed. Bottom right: Correlation

map illustrating the correlation coefficient  $r$  of word-reading speed (number of correctly read words per minute) in second grade, with the mean N1 amplitude at the baseline test time in kindergarten at every channel. The strong positive correlation over posterior (right occipital, parietal, and temporal channels) indicates that an increase in the print-sensitive N1 negativity is associated with poorer reading skills (fewer words read). Note, all anterior correlations are negative.

## Tables

**Table 1**

*Behavioural Assessment and Demographic Data of Future Poor and Regular Readers at T0.*

Longitudinal GG Training Sample					
Kindergarten	Regular (RR), n=13		Poor (PR), n=12		<i>t</i> -Test
	M	SD	M	SD	<i>p</i>
Gender (Male : Female)	6:7		6:6		
Age (years)	6.4	.3	6.3	.3	0.47
Estimated Nonverbal IQ (RAVEN)	117.9	19.4	108.9	14.3	0.20
Estimated SES Parents	16.8	3.2	16.4	3.2	0.74
ARHQ	.3	.1	.3	.1	0.98
<b>Letter LC *</b>	<b>9.7</b>	<b>7.6</b>	<b>3.8</b>	<b>3.0</b>	<b>0.019</b>
Letter UC	12.9	9.0	9.0	5.9	0.21
Read (No. Words)	.9	1.7	.2	.6	0.15
<b>Receptive Vocabulary (MSVK) [pc] **</b>	<b>78.2</b>	<b>18.4</b>	<b>52.2</b>	<b>23.4</b>	<b>0.005</b>
Word Meaning (MSVK) [pc]	43.9	28.4	48.3	21.5	0.67
RAN (*)	42.0	7.3	46.8	4.8	0.062
BISC Risk Score	1.5	1.9	1.6	1.3	0.95
BISC PAns	18.5	2.3	17.9	2.4	0.57
<b>BISC PAbs*</b>	<b>18.0</b>	<b>1.5</b>	<b>15.8</b>	<b>2.8</b>	<b>0.020</b>
BISC Pseudoword-Repetition	5.8	2.4	6.2	2.2	0.73
BISC Visual Word Comparison,Accuracy	10.8	1.2	11.0	1.2	0.63

BISC Visual Word Comparison, Time <sup>(*)</sup>	4.4	2.8	6.5	3.4	0.093
BISC Colour Naming	7.8	.6	8.0	.0	0.19
<b>BISC RAN-Colours, con *</b>	<b>7.2</b>	<b>.6</b>	<b>6.1</b>	<b>1.6</b>	<b>0.032</b>
BISC RAN-Colours, inc	9.1	3.3	8.6	3.5	0.72
<b>BISC Rhyme *</b>	<b>9.8</b>	<b>.4</b>	<b>8.8</b>	<b>1.4</b>	<b>0.033</b>
<b>BISC Phoneme Association *</b>	<b>9.0</b>	<b>1.8</b>	<b>9.5</b>	<b>.7</b>	<b>0.037</b>
BISC Syllable Segmentation <sup>(*)</sup>	8.2	1.3	7.0	2.0	0.083
BISC Phoneme Extraction	9.5	0.9	8.4	2.2	0.15

Abbreviations: SES, estimated socio-economic status; LC, lower case; UC, upper case; ARHQ, adult reading history questionnaire; MSVK, "Marburger Sprachverständnistest für Kinder", test battery to examine word meaning and receptive vocabulary; BISC, "Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten BISC", literacy screening battery that tests phonological awareness as well as memory and attentional skills; PAns, phonological awareness in a narrow sense, PAbs, phonological awareness in a broad sense; RAN, rapid automatized naming; inc, incongruent; con, congruent; [pc] percentile. \*\* $p < 0.01$ , \* $p < 0.05$ , <sup>(\*)</sup> $p < 0.1$ .

**Table 2***Behavioural Assessment and Demographic Data of Poor and Regular Readers of the Longitudinal GG and the Control Cohorts in Second Grade*

2nd Grade	All					Longitudinal GG Training Cohort (GG)					Control Cohort (Ctrl)				
	Regular (RR),		Poor (PR),		<i>t</i> -Test	Regular (RR),		Poor (PR),		<i>t</i> -Test	Regular (RR),		Poor (PR),		<i>t</i> -Test
Groups	n=23		n=23			n=13		n=12			n=10		n=11		
Gender (Male : Female)	10:13		12:11			6:7		6:6			4:6		6:5		
	M	SD	M	SD	<i>p</i>	M	SD	M	SD	<i>p</i>	M	SD	M	SD	<i>p</i>
Age (years)	8.3	.4	8.3	.3	0.89	8.4	.3	8.4	.3	0.95	8.3	.5	8.2	.3	0.96
<b>Word/text reading (SLRT) [pc]</b>	<b>63.0</b>	<b>16.5</b>	<b>8.3</b>	<b>7.3</b>	<b>&lt;0.001***</b>	<b>60.8</b>	<b>13.4</b>	<b>9.6</b>	<b>6.2</b>	<b>&lt;0.001***</b>	<b>66.0</b>	<b>20.3</b>	<b>7.0</b>	<b>8.3</b>	<b>&lt;0.001***</b>
> No. Correctly Read Words Per Minute	85.4	16.9	30.9	12.1	<0.001***	82.1	13.8	33.7	11.1	<0.001***	89.7	20.2	27.9	13.0	<0.001***
<b>Word Reading (SLRT-II) [pc]</b>	<b>58.6</b>	<b>22.9</b>	<b>14.9</b>	<b>10.8</b>	<b>&lt;0.001***</b>	<b>55.8</b>	<b>23.1</b>	<b>19.3</b>	<b>10.4</b>	<b>&lt;0.001***</b>	<b>62.4</b>	<b>23.4</b>	<b>10.2</b>	<b>9.4</b>	<b>&lt;0.001***</b>
> No. Correctly Read Words Per Minute	49.4	13.0	23.3	7.9	<0.001***	46.6	11.9	26.3	7.6	<0.001***	52.9	14.1	19.9	7.1	<0.001***
<b>Pseudoword Reading (SLRT-II) [pc]</b>	<b>48.3</b>	<b>24.9</b>	<b>14.8</b>	<b>11.3</b>	<b>&lt;0.001***</b>	<b>46.7</b>	<b>25.9</b>	<b>17.7</b>	<b>10.9</b>	<b>0.002**</b>	<b>50.4</b>	<b>24.8</b>	<b>11.7</b>	<b>11.4</b>	<b>&lt;0.001***</b>
> No. Correctly Read Pseudowords Per Minute	30.7	8.7	19.2	5.3	<0.001***	29.6	8.1	21.0	4.3	0.003	32.0	9.7	17.2	5.7	<0.001***
<b>Spelling (SLRT) [pc]</b>	<b>59.4</b>	<b>27.6</b>	<b>22.2</b>	<b>19.0</b>	<b>&lt;0.001***</b>	<b>66.1</b>	<b>25.4</b>	<b>29.7</b>	<b>22.6</b>	<b>&lt;0.001***</b>	<b>50.7</b>	<b>29.2</b>	<b>14.1</b>	<b>9.6</b>	<b>0.003**</b>
<b>BAKO Total Phonological Score [pc]</b>	<b>50.3</b>	<b>20.7</b>	<b>33.7</b>	<b>15.9</b>	<b>0.004**</b>	48.5	22.5	36.3	15.8	0.13	<b>52.7</b>	<b>19.1</b>	<b>31.0</b>	<b>16.3</b>	<b>0.011*</b>
> Pseudoword Segmentation [pc]	53.8	26.6	45.3	27.3	0.29	59.7	25.2	56.0	22.9	0.71	46.2	27.8	33.7	27.9	0.32

> Vowel Substitution [pc]	45.0	19.3	39.2	18.5	0.31	38.8	15.9	34.8	15.9	0.53	53.0	21.1	44.0	20.7	0.34
> Rest Word Identification [pc]	43.4	20.2	36.1	22.1	0.25	36.9	19.0	34.8	23.6	0.80	51.8	19.4	37.5	21.4	0.13
> <b>Phoneme Inversion [pc]</b>	<b>59.1</b>	<b>24.0</b>	<b>32.4</b>	<b>19.7</b>	<b>&lt;0.001***</b>	<b>59.5</b>	<b>25.3</b>	<b>35.8</b>	<b>22.7</b>	<b>0.022*</b>	<b>58.7</b>	<b>23.5</b>	<b>28.6</b>	<b>16.1</b>	<b>0.003**</b>
> <b>Sound Categorization [pc]</b>	<b>55.6</b>	<b>31.0</b>	<b>27.9</b>	<b>22.4</b>	<b>0.001**</b>	<b>53.6</b>	<b>30.5</b>	<b>26.5</b>	<b>18.0</b>	<b>0.013*</b>	<b>58.2</b>	<b>33.2</b>	<b>29.5</b>	<b>27.3</b>	<b>0.042*</b>
> Vowel Length Determination [pc]	61.8	25.5	47.4	28.5	0.079 <sup>(*)</sup>	65.5	19.8	55.8	28.9	0.34	57.0	32.0	38.4	26.4	0.16
> Word Inversion [pc]	43.7	27.7	35.3	25.8	0.29	40.5	27.8	36.9	23.6	0.74	48.0	28.4	33.5	29.2	0.27
<b>HAWIK Estimated Verbal IQ</b>	<b>125.9</b>	<b>12.2</b>	<b>113.5</b>	<b>12.7</b>	<b>0.002**</b>	<b>126.5</b>	<b>12.0</b>	<b>114.6</b>	<b>14.2</b>	<b>0.032*</b>	<b>125.0</b>	<b>13.1</b>	<b>112.3</b>	<b>11.3</b>	<b>0.027*</b>
HAWIK Estimated Nonverbal IQ	113.9	10.1	109.3	12.7	0.19	114.2	11.9	109.6	15.4	0.41	113.5	7.8	109.1	9.7	0.27
RAN	<b>29.5</b>	<b>5.4</b>	<b>33.7</b>	<b>4.3</b>	<b>0.005**</b>	<b>28.7</b>	<b>4.9</b>	<b>32.7</b>	<b>4.5</b>	<b>0.048*</b>	30.4	6.0	34.8	3.8	0.062 <sup>(*)</sup>
Letters UC	28.7	.5	28.3	.8	0.055 <sup>(*)</sup>	28.6	.5	28.3	.8	0.17	28.8	.4	28.4	.9	0.19
<b>Letters LC</b>	<b>28.4</b>	<b>.9</b>	<b>27.1</b>	<b>1.5</b>	<b>&lt;0.001***</b>	<b>28.5</b>	<b>.7</b>	<b>27.5</b>	<b>1.2</b>	<b>0.020*</b>	<b>28.3</b>	<b>1.2</b>	<b>26.7</b>	<b>1.7</b>	<b>0.023*</b>
WM Forward <sup>a</sup>	6.3	1.5	5.8	1.3	0.25	6.2	1.3	6.1	1.3	0.90	6.5	1.8	5.4	1.2	0.16
WM Backward <sup>a</sup>	4.5	1.4	4.0	1.2	0.29	4.2	1.3	4.3	1.2	0.73	4.9	1.4	3.7	1.2	0.063 <sup>(*)</sup>

For the three core reading measures, percentiles [pc] and raw scores are tabulated. Abbreviations: SLRT, “Salzburger Lese- und Rechtschreibtest”, test battery examining reading and spelling; BAKO, “Basiskompetenzen für Lese-Rechtschreibleistungen”, test battery to test reading and spelling competences at school; RAN, rapid automatized naming; LC, lower case; UC, upper case; WM, working memory. <sup>(a)</sup> Data available for 9 poor readers (ctrl group) only. \*\*\*  $p < 0.001$ . \*\*  $p < 0.01$ , \*  $p < 0.05$ , <sup>(\*)</sup>  $p < 0.1$ .

**Table 3***Multiple Regression Analyses to Predict Word Reading Speed in Second Grade*

<b>Analysis 1) Behavioural and Demographic Measures</b>				
	2 <sup>nd</sup> Grade Word Reading	B	SE B	$\beta$
Step 1	Constant	71.28	12.62	
	RAN	-0.78	0.28	-0.46**
Step 2	Constant	38.23	19.77	
	RAN	-0.71	0.27	-0.42*
	IQ (RAVEN)	0.27	0.13	0.33*
Average $R^2=.21$ ( $p=0.010$ ) for step 1; $\Delta R^2=.11$ ( $p=0.045$ ) for step 2				
<b>Analysis 2) Electrophysiological Measures</b>				
	2 <sup>nd</sup> Grade Word Reading	B	SE B	$\beta$
Step 1	Constant	39.36	2.21	
	N1 ROT Diff. W-FF	1.58	0.51	0.50**
<b>Analysis 3) Combined Analysis of Behavioural and Electrophysiological Measures</b>				
	2 <sup>nd</sup> Grade Word Reading	B	SE B	$\beta$
Step 1	Constant	38.23	19.77	
	RAN	-0.71	0.27	-0.42*
	IQ (RAVEN)	0.27	0.13	0.33*
Step 2	Constant	41.28	17.9	
	RAN	-0.61	0.25	-0.36*
	IQ (RAVEN)	0.22	0.12	0.27 <sup>(*)</sup>
	N1 ROT Diff. W-FF	1.25	0.46	0.39*
Average $R^2=.32$ ( $p=0.005$ ) for step 1; $\Delta R^2=.15$ ( $p=0.012$ ) for step 2				

Abbreviations: RAN, rapid automatized naming; W, words; FF, falsefonts; ROT, right occipito-temporal sites; Diff., difference. <sup>(\*)</sup>  $p<0.1$ , \*  $p<0.05$ , \*\*  $p<0.01$ .

## References

- Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cereb Cortex*, 4(5), 544-554.
- Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb Cortex*, 9(5), 415-430.
- Araujo, S., Bramao, I., Faisca, L., Petersson, K. M., & Reis, A. (2012). Electrophysiological correlates of impaired reading in dyslexic pre-adolescent children. *Brain Cogn*, 79(2), 79-88. doi: S0278-2626(12)00033-4 [pii] 10.1016/j.bandc.2012.02.010
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (pp. CD-ROM). Philadelphia, University of Pennsylvania: Linguistic Data Consortium.
- Bach, S., Brandeis, D., Hofstetter, C., Martin, E., Richardson, U., & Brem, S. (2010). Early emergence of deviant frontal fMRI activity for phonological processes in poor beginning readers. *Neuroimage*, 53(1), 682-693. doi:10.1016/j.neuroimage.2010.06.039
- Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage*, 82, 605-615. doi: 10.1016/j.neuroimage.2013.05.062
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc Natl Acad Sci U S A*, 104(21), 9087-9092. doi: 0703300104 [pii] 10.1073/pnas.0703300104
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *J Cogn Neurosci*, 8(6), 551-565. doi: 10.1162/jocn.1996.8.6.551
- Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: time course and scalp distribution. *Journal of Cognitive Neuroscience*, 11(3), 235-260.
- Bowey, J. A. (1995). Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. *Journal of Educational Psychology*, 87, 476-487.

- Brem, S., Bach, S., Kucian, K., Guttorm, T. K., Martin, E., Lyytinen, H. (2010). Brain sensitivity to print emerges when children learn letter-speech sound correspondences. *Proc Natl Acad Sci U S A*, 107(17), 7939-7944. doi: 0904402107 [pii] 10.1073/pnas.0904402107
- Brem, S., Bucher, K., Halder, P., Summers, P., Dietrich, T., Martin, E. (2006). Evidence for developmental changes in the visual word processing network beyond adolescence. *Neuroimage*, 29(3), 822-837.
- Brem, S., Halder, P., Bucher, K., Summers, P., Martin, E., & Brandeis, D. (2009). Tuning of the visual word processing system: distinct developmental ERP and fMRI effects. *Hum Brain Mapp*, 30(6), 1833-1844. doi: 10.1002/hbm.20751
- Brem, S., Lang-Dullenkopf, A., Maurer, U., Halder, P., Bucher, K., & Brandeis, D. (2005). Neurophysiological signs of rapidly emerging visual expertise for symbol strings. *Neuroreport*, 16(1), 45-48.
- Catts, H. W., Fey, M. E., Zhang, X. L., & Tomblin, J. B. (2001). Estimating the Risk of Future Reading Difficulties in Kindergarten Children *Language, Speech, and Hearing Services in Schools* 32, 38-50.
- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(Pt 5), 1054-1069.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychol Rev*, 108(1), 204-256.
- de Jong, P., & van der Leij, A. (1999). Specific Contributions of Phonological Abilities to Early Reading Acquisition: Results From a Dutch Latent Variable Longitudinal Study. *Journal of Educational Psychology*, 91(3), 450-476.
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384-398.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends Cogn Sci*, 9(7), 335-341.
- Elben, C. E., & Lohaus, A. (2000). *Marburger Sprachverständnistest für Kinder*. Göttingen: Hogrefe-Verlag.
- Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nat Neurosci*, 6(4), 428-432.



- Goswami, U., Wang, H. L., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *J Cogn Neurosci*, 23(2), 325-337. doi: 10.1162/jocn.2010.21453
- Guttorm, T. K., Leppanen, P. H., Hamalainen, J. A., Eklund, K. M., & Lyytinen, H. J. (2010). Newborn event-related potentials predict poorer pre-reading skills in children at risk for dyslexia. *J Learn Disabil*, 43(5), 391-401. doi: 0022219409345005 [pii] 10.1177/0022219409345005
- Guttorm, T. K., Leppanen, P. H., Poikkeus, A. M., Eklund, K. M., Lyytinen, P., & Lyytinen, H. (2005). Brain event-related potentials (ERPs) measured at birth predict later language development in children with and without familial risk for dyslexia. *Cortex*, 41(3), 291-303.
- Helenius, P., Tarkiainen, A., Cornelissen, P., Hansen, P. C., & Salmelin, R. (1999). Dissociation of normal feature analysis and deficient processing of letter-strings in dyslexic adults. *Cerebral Cortex*, 9(5), 476-483.
- Hoeft, F., McCandliss, B. D., Black, J. M., Gantman, A., Zakerani, N., Hulme, C. (2011). Neural systems predicting long-term outcome in dyslexia. *Proc Natl Acad Sci U S A*, 108(1), 361-366. doi: 1008950108 [pii] 10.1073/pnas.1008950108
- Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behav Neurosci*, 121(3), 602-613. doi: 2007-08636-017 [pii] 10.1037/0735-7044.121.3.602
- Jansen, A., Mannhaupt, G., Marx, H., & Skowronek, H. (1999). *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC)*. Göttingen: Hogrefe-Verlag GmbH.
- Juel, C. (1986). Learning to Read and Write: A Longitudinal Study of 54 Children From First Through Fourth Grades. *Journal of Educational Psychology*, 80(4), 437-447.
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol*, 111(10), 1745-1758. doi: S1388-2457(00)00386-2 [pii]
- Kast, M., Elmer, S., Jancke, L., & Meyer, M. (2010). ERP differences of pre-lexical processing between dyslexic and non-dyslexic children. *Int J Psychophysiol*, 77(1), 59-69. doi: S0167-8760(10)00099-1 [pii] 10.1016/j.ijpsycho.2010.04.003
- Landerl, K., Wimmer, H., & Moser, E. (1997). *Der Salzburger Lese- und Rechtschreibtest (SLRT)*. Bern: Huber Verlag.

- Kronschabel, J., Schmid, R., Maurer, U., Brandeis, D. (2013). Visual print tuning deficits in dyslexic adolescents under minimized phonological demands. *Neuroimage*, 74(1), 58-69.
- Lefly, D. L., & Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *J Learn Disabil*, 33(3), 286-296.
- Lehmann, D. (1984). EEG assessment of brain activity: spatial aspects, segmentation and imaging. *Int J Psychophysiol*, 1(3), 267-276.
- Leppanen, P. H., Hamalainen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex*, 46(10), 1362-1376. doi: S0010-9452(10)00175-9 [pii] 10.1016/j.cortex.2010.06.003
- Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M. L. (2004). The development of children at familial risk for dyslexia: birth to early school age. *Ann Dyslexia*, 54(2), 184-220.
- Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: a learning tool for reading acquisition. *Scand J Psychol*, 50(6), 668-675. doi: SJOP791 [pii] 10.1111/j.1467-9450.2009.00791.x
- Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A. M., & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology*, 59(2), 109-126.
- Mahé, G., Bonnefond, A., Gavens, N., Dufour, A., Doignon-Camus, A. (2012). Impaired visual expertise for print in French adults with dyslexia as shown by N170 tuning. *Neuropsychologia*, 50, 3200-3206.
- Maurer, U., Blau, V. C., Yoncheva, Y. N., & McCandliss, B. D. (2010). Development of visual expertise for reading: rapid emergence of visual familiarity for an artificial script. *Dev Neuropsychol*, 35(4), 404-422. doi: 923935370 [pii] 10.1080/87565641.2010.480916
- Maurer, U., Brem, S., Bucher, K., & Brandeis, D. (2005). Emerging neurophysiological specialization for letter strings. *J Cogn Neurosci*, 17(10), 1532-1552.
- Maurer, U., Brem, S., Bucher, K., Kranz, F., Benz, R., Steinhausen, H. C. (2007). Impaired tuning of a fast occipito-temporal response for print in dyslexic children learning to read. *Brain*, 130(Pt 12), 3200-3210.

- Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P. (2006). Coarse neural tuning for print peaks when children learn to read. *Neuroimage*, 33(2), 749-758.
- Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biol Psychiatry*, 66(4), 341-348. doi: S0006-3223(09)00298-4 [pii] 10.1016/j.biopsych.2009.02.031
- Maurer, U., & McCandliss, B. D. (2007). The development of visual expertise for words: the contribution of electrophysiology. In Elena L. Grigorenko and Adam J. Naples (Eds.), *Single-word reading: biological and behavioural perspectives* (pp. 43-64). Mahwah, NJ: Lawrence Erlbaum Associates.
- Maurer, U., Schulz, E., Brem, S., der Mark, S., Bucher, K., Martin, E. (2011). The development of print tuning in children with dyslexia: Evidence from longitudinal ERP data supported by fMRI. *Neuroimage*, 57(3), 714-722. doi: S1053-8119(10)01348-0 [pii] 10.1016/j.neuroimage.2010.10.055
- Molfese, D. L. (2000). Predicting dyslexia at 8 years of age using neonatal brain responses. *Brain Lang*, 72(3), 238-245. doi: 10.1006/brln.2000.2287 S0093-934X(00)92287-9 [pii]
- Moll, K., & Landerl, K. (2010). *Lese- und Rechtschreibtest (SLRT-II) Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)*. Bern: Verlag Hans Huber.
- Mugnaini, D., Lassi, S., La Malfa, G., Albertini, G. (2009). Internalizing correlates of dyslexia. *World Journal of Pediatrics*, 5(4), 255-264.
- Parviainen, T., Helenius, P., Poskiparta, E., Niemi, P., & Salmelin, R. (2006). Cortical sequence of word perception in beginning readers. *J Neurosci*, 26(22), 6052-6061.
- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci*, 15(6), 246-253. doi: S1364-6613(11)00057-X [pii] 10.1016/j.tics.2011.04.001
- Proverbio, A. M., Zani, A., & Adorni, R. (2008). The left fusiform area is affected by written frequency of words. *Neuropsychologia*, 46(9), 2292-2299. doi: S0028-3932(08)00117-6 [pii] 10.1016/j.neuropsychologia.2008.03.024
- Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppanen, P. H., Poikkeus, A. M. (2007). Very early phonological and language skills: estimating individual risk of reading disability. *J Child Psychol Psychiatry*, 48(9), 923-931. doi: JCPP1763 [pii] 10.1111/j.1469-7610.2007.01763.x

- Räsänen, P., Salminen, J., Wilson, a. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, 24(4), 450-472.
- Raschle, N. M., Chang, M., & Gaab, N. (2010). Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage*. doi: S1053-8119(10)01255-3 [pii] 10.1016/j.neuroimage.2010.09.055
- Raschle, N. M., Zuk, J., & Gaab, N. (2012). Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. *Proc Natl Acad Sci U S A*, 109(6), 2156-2161. doi: 1107721109 [pii] 10.1073/pnas.1107721109
- Raven, J. C. (2002). *Coloured Progressive Matrices (CPM)* ((3., neu norm. Aufl.) ed.). Frankfurt: Harcourt Test Services.
- Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage*, 20(3), 1609-1624.
- Saine, N. L., Lerkkanen, M. K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. *Child Dev*, 82(3), 1013-1028. doi: 10.1111/j.1467-8624.2011.01580.x
- Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Dev*, 61(6), 1728-1743.
- Schendan, H. E., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual categorization of words and faces within 150 ms. *Psychophysiology*, 35(3), 240-251.
- Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *J Cogn Neurosci*, 18(9), 1453-1465. doi: 10.1162/jocn.2006.18.9.1453
- Stock, C., Marx, P., & Schneider, W. (2003). BAKO1-4. Basiskompetenzen für Lese-Rechtschreibleistungen. Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis zum vierten Grundschuljahr. Göttingen: Beltz Test GmbH.
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychol Sci*, 12(1), 43-47.
- Tewes, U., Rossmann, P., & Schallberger, U. (1999). *Hamburg-Wechsler-Intelligenz-Test für Kinder III (HAWIK III). Handbuch und Testanweisung*. Bern: Huber.

- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143-156.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and the acquisition of reading skills. *Psychol. Bull.*, 101 192–212.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology*, 33, 468-479.
- Willcutt, E. G. & Pennington B. F. Psychiatric comorbidity in children and adolescents with reading disability. *J Child Psychol Psychiatry* 2000;41:1039-1048.
- Wimmer, H., Landerl, K., Linortner, R., & Hummer, P. (1991). The relationship of phonemic awareness to reading acquisition: more consequence than precondition but still important. *Cognition*, 40(3), 219-249.
- Wimmer, H., Mayringer, H., & Landerl, K. (2000). The Double-Deficit Hypothesis and Difficulties in Learning to Read a Regular Orthography *Journal of Educational Psychology*, 92(4), 668-680.
- Wolf, M. (1986). Automaticity, retrieval processes, and reading: Longitudinal study in average and impaired readers. *Child Dev*, 57, 988-1005.
- Yoncheva, Y. N., Blau, V. C., Maurer, U., & McCandliss, B. D. (2010). Attentional focus during learning impacts N170 ERP responses to an artificial script. *Dev Neuropsychol*, 35(4), 423-445. doi: 923934384 [pii] 10.1080/87565641.2010.480918
- Zhang, M., Jiang, T., Mei, L., Yang, H., Chen, C., Xue, G. (2011). It's a word: early electrophysiological response to the character likeness of pictographs. *Psychophysiology*, 48(7), 950-959. doi: 10.1111/j.1469-8986.2010.01153.